GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2004, 01:42:25 ; Search time 137 Seconds
(without alignments)
1492.378 Million cell updates/sec

Title: US-09-805-020-72
Perfect score: 3506
Sequence: 1 MSPPFLRIGLSNFDCGSCQSC.....LVSFLFLLVSNLHVANNDYY 648

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2844	81.1	571	11	Q8CAV6	Q8cav6 mus musculu
2	2293	65.4	469	11	Q8C3L7	Q8c3l7 mus musculu
3	2120	60.5	683	13	Q7SZH8	Q7szh8 xenopus lae
4	2100	59.9	683	13	Q7SZH7	Q7szh7 xenopus lae
5	2045.5	58.3	676	4	Q86XJ6	Q86xj6 homo sapien
6	2018.5	57.6	684	13	Q7ZUC5	Q7zuc5 brachydanio
7	1757	50.1	464	11	Q920N8	Q920n8 mus musculu
8	1566.5	44.7	487	11	Q80UN7	Q80un7 mus musculu
9	1491	42.5	724	5	O61225	O61225 scypha raph
10	1447	41.3	763	5	O62569	O62569 suberites d
11	1345	38.4	754	5	Q8MXB6	Q8mxb6 limulus pol
12	1290.5	36.8	707	5	Q20953	Q20953 caenorhabdi
13	1247.5	35.6	683	11	Q8K2K8	Q8k2k8 mus musculu
14	1241.5	35.4	683	4	Q8NE03	Q8ne03 homo sapien
15	1235	35.2	661	5	O01669	O01669 hydra atten
16	1166.5	33.3	670	13	Q7SY24	Q7sy24 brachydanio
17	1160	33.1	670	13	Q8JFZ9	Q8jfz9 fugu rubrip
18	1145	32.7	756	5	O61224	O61224 scypha raph
19	1144.5	32.6	680	5	P90980	P90980 caenorhabdi
20	1144.5	32.6	680	5	Q8MQ88	Q8mq88 caenorhabdi
21	1144.5	32.6	682	5	Q8MQ87	Q8mq87 caenorhabdi
22	1144.5	32.6	682	5	P90981	P90981 caenorhabdi
23	1144.5	32.6	936	5	Q19024	Q19024 caenorhabdi
24	1129	32.2	668	13	Q7T2C5	Q7t2c5 brachydanio
25	1097	31.3	670	5	O01715	O01715 hydra atten
26	1076.5	30.7	674	5	O01716	O01716 hydra atten
27	1070.5	30.5	554	5	Q95T78	Q95t78 drosophila
28	1067	30.4	673	5	O62567	O62567 suberites d
29	998	28.5	677	5	O96997	O96997 geodia cydo
30	975	27.8	685	5	O76850	O76850 calliphora
31	949.5	27.1	1035	3	Q862V2	Q86zv2 pichia past
32	892	25.4	1161	3	Q8J213	Q8j213 kluyveromyc
33	885	25.2	1157	3	Q9HF10	Q9hf10 blumeria gr
34	883	25.2	1194	3	Q9Y792	Q9y792 sporothrix
35	871.5	24.9	1170	3	Q9UVJ5	Q9uvj5 botrytis ci
36	860.5	24.5	1185	3	Q873Y9	Q873y9 leptosphaer
37	857	24.4	1136	3	Q9HGK8	Q9hgk8 tuber borch
38	853	24.3	1182	3	Q9Y7C1	Q9y7c1 magnaporthe
39	850	24.2	697	5	O96942	O96942 rhabdocalyp
40	848.5	24.2	447	5	Q86M17	Q86m17 branchiosto
41	845	24.1	991	3	Q96VF6	Q96vf6 tuber magna
42	775	22.1	442	13	Q801L5	Q80115 scyliorhinu
43	739.5	21.1	606	5	Q9V782	Q9v782 drosophila
44	739.5	21.1	606	5	Q9GSZ3	Q9gsz3 drosophila
45	739.5	21.1	606	5	Q8MT38	Q8mt38 drosophila

ALIGNMENTS

RESULT 1
Q8CAV6
ID Q8CAV6 PRELIMINARY; PRT; 571 AA.
AC Q8CAV6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Protein kinase C.
GN A130035A12RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK037664; BAC29843.1; -.
DR MGD; MGI:2442369; A130035A12Rik.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00130; DAG_PE-bind; 2.

```
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00008; DAGPEDOMAIN.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00109; C1; 2.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 2.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ   SEQUENCE   571 AA;  64932 MW;  710ED91673EA3C59 CRC64;

  Query Match             81.1%;  Score 2844;  DB 11;  Length 571;
  Best Local Similarity   94.5%;  Pred. No. 7e-246;
  Matches  520;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MSPFLRIGLSNFDCGSCQSCQGEAVNPYCAVLVKEYVESENGQMYIQKKPTMYPPWDSTF 60
              ||||||||||||||:||:||||||||||||||||||||||||||||||||||||||||||
Db          1 MSPFLRIGLSNFDCGTCQACQGEAVNPYCAVLVKEYVESENGQMYIQKKPTMYPPWDSTF 60

Qy         61 DAHINKGRVMQIIVKGKNVDLISETTVELYSLAERCRKNNGKTEIWLELKPQGRMLMNAR 120
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db         61 DAHINKGRVMQIIVKGKNVDLISETTVELYSLAERCRKNNGRTEIWLELKPQGRMLMNAR 120

Qy        121 YFLEMSDTKDMNEFETEGFFALHQRRGAIKQAKVHHVKCHEFTATFFPQPTFCSVCHEFV 180
              ||||||||||:||| ||||||||||||||||||||||||||||||||||||||||||||
Db        121 YFLEMSDTKDMSEFENEGFFALHQRRGAIKQAKVHHVKCHEFTATFFPQPTFCSVCHEFV 180

Qy        181 WGLNKQGYQCRQCNAAIHKKCIDKVIAKCTGSAINSRETMFHKERFKIDMPHRFKVYNYK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 WGLNKQGYQCRQCNAAIHKKCIDKVIAKCTGSAINSRETMFHKERFKIDMPHRFKVYNYK 240

Qy        241 SPTFCEHCGTLLWGLARQGLKCDACGMNVHHRCQTKVANLCGINQKLMAEALAMIESTQQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SPTFCEHCGTLLWGLARQGLKCDACGMNVHHRCQTKVANLCGINQKLMAEALAMIESTQQ 300

Qy        301 ARCLRDTEQIFREGPVEIGLPCSIKNEARPPCLPTPGKREPQGISWESPLDEVDKMCHLP 360
              || |||:| |||||||||||||| ||| ||||:|||||||||||:|||| :|     |
Db        301 ARSLRDSEHIFREGPVEIGLPCSTKNETRPPCVPTPGKREPQGISWDSPLDGSNKSAGPP 360

Qy        361 EPELNKERPSLQIKLKIEDFILHKMLGKGSFGKVFLAEFKKTNQFFAIKALKKDVVLMDD 420
              |||::  | |||:|||||:|||||||||||||||||:|||||||||||||||||||||
Db        361 EPEVSMRRTSLQLKLKIDDFILHKMLGKGSFGKVFLAEFKRTNQFFAIKALKKDVVLMDD 420

Qy        421 DVECTMVEKRVLSLAWEHPFLTHMFCTFQTKENLFFVMEYLNGGDLMYHIQSCHKFDLSR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 DVECTMVEKRVLSLAWEHPFLTHMFCTFQTKENLFFVMEYLNGGDLMYHIQSCHKFDLSR 480

Qy        481 ATFYAAEIILGLQFLHSKGIVYRDLKLDNILLDKDGHIKIADFGMCKENMLGDAKTNTFC 540
              |||||||:|||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        481 ATFYAAEVILGLQFLHSKGIVYRDLKLDNILLDRDGHIKIADFGMCKENMLGDAKTNTFC 540

Qy        541 GTPDYIAPEI 550
              |||||||||:
Db        541 GTPDYIAPEV 550


RESULT 2
Q8C3L7
ID   Q8C3L7      PRELIMINARY;      PRT;   469 AA.
AC   Q8C3L7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Protein kinase C.
GN   A130035A12RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK085546; BAC39469.1; -.
DR   MGD; MGI:2442369; A130035A12Rik.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0005489; F:electron transporter activity; IEA.
DR   GO; GO:0004672; F:protein kinase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR008973; C2_CaLB.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000719; Prot_kinase.
DR   Pfam; PF00130; DAG_PE-bind; 2.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00008; DAGPEDOMAIN.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00109; C1; 2.
DR   PROSITE; PS00190; CYTOCHROME_C; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 2.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
SQ   SEQUENCE   469 AA;  53543 MW;  6F3D50C4DC2DE1A2 CRC64;

  Query Match             65.4%;  Score 2293;  DB 11;  Length 469;
  Best Local Similarity   90.5%;  Pred. No. 1.3e-196;
  Matches  419;  Conservative   22;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 MSPFLRIGLSNFDCGSCQSCQGEAVNPYCAVLVKEYVESENGQMYIQKKPTMYPPWDSTF 60
              ||||||||||||||:||:||||||||||||||||||||||||||||:||||||||||||
Db          1 MSPFLRIGLSNFDCGTCQACQGEAVNPYCAVLVKEYVESENGQMYIQEKPTMYPPWDSTF 60

Qy         61 DAHINKGRVMQIIVKGKNVDLISETTVELYSLAERCRKNNGKTEIWLELKPQGRMLMNAR 120
              ||||||||||||||:||||||||||||||||||:|:||||||||||||||:| |
Db         61 DAHINKGRVMQIIVKGKDVDLISETTVELYSLAERCRKNHGRTEIWLELKPQGRMLINPR 120

Qy        121 YFLEMSDTKDMNEFETEGFFALHQRRGAIKQAKVHHVKCHEFTATFFPQPTFCSVCHEFV 180
              ||||||||||:||| |||| ||||||||||||||||||||||||||||||||||||||
Db        121 YFLEMSDTKDMSEFENEGFFPLHQRRGAIKQAKVHHVKCHEFTATFFPQPTFCSVCHEFV 180

Qy        181 WGLNKQGYQCRQCNAAIHKKCIDKVIAKCTGSAINSRETMFHKERFKIDMPHRFKVYNYK 240
              ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        181 WGLNKQGYQCRQCNAAIHKKCIDKVIAKCTGSAIHSRETMFHKERFKIDMPHRFKVYNYK 240

Qy        241 SPTFCEHCGTLLWGLARQGLKCDACGMNVHHRCQTKVANLCGINQKLMAEALAMIESTQQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SPTFCEHCGTLLWGLARQGLKCDACGMNVHHRCQTKVANLCGINQKLMAEALAMIESTQQ 300

Qy        301 ARCLRDTEQIFREGPVEIGLPCSIKNEARPPCLPTPGKREPQGISWESPLDEVDKMCHLP 360
              || |||:| |||||||||||||| ||| ||||:|||||||||||:|||| :|     |
Db        301 ARSLRDSEHIFREGPVEIGLPCSTKNETRPPCVPTPGKREPQGISWDSPLDGSNKSAGPP 360

Qy        361 EPELNKERPSLQIKLKIEDFILHKMLGKGSFGKVFLAEFKKTNQFFAIKALKKDVVLMDD 420
              |||::  | |||:|||||:|||||||||||||||||:|||||||||||||||||||||
Db        361 EPEVSMRRTSLQLKLKIDDFILHKMLGKGSFGKVFLAEFKRTNQFFAIKALKKDVVLMDD 420

Qy        421 DVECTMVEKRVLSLAWEHPFLTHMFCTFQTKENLFFVMEYLNG 463
              |||||||||||||||||||||||||||||  : : : :| |
Db        421 DVECTMVEKRVLSLAWEHPFLTHMFCTFQTKVCVCYSLSFLFG 463


RESULT 3
Q7SZH8
ID   Q7SZH8      PRELIMINARY;      PRT;   683 AA.
AC   Q7SZH8;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 23:31:44 ; Search time 140 Seconds
(without alignments)
1307.791 Million cell updates/sec

Title: US-09-805-020-72
Perfect score: 3506
Sequence: 1 MSPFLRIGLSNFDCGSCQSC.....LVSFLFLLVSNLHVANNDYY 648

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3506	100.0	648	5	ABG79705	Abg79705 Tumour in
2	3344	95.4	615	5	ABG79678	Abg79678 Tumour in
3	3337	95.2	706	4	AAG62870	Aag62870 Amino aci
4	3337	95.2	706	6	ABP56907	Abp56907 Human pro
5	3327	94.9	706	4	AAB60648	Aab60648 Human pro
6	3327	94.9	706	6	ABR63656	Abr63656 Human pro
7	2057.5	58.7	673	7	ADE57523	Ade57523 Rat Prote
8	2057.5	58.7	673	7	ADE57525	Ade57525 Rat Prote
9	2037.5	58.1	676	4	AAB48037	Aab48037 Human pro
10	1499	42.8	704	3	AAY91091	Aay91091 Caenorhab
11	1499	42.8	704	3	AAY59506	Aay59506 C. elegan
12	1343	38.3	567	3	AAY59507	Aay59507 C. elegan
13	1343	38.3	567	3	AAB03667	Aab03667 Nematode
14	1298.5	37.0	737	5	AAO18490	Aao18490 Human ins
15	1298.5	37.0	737	7	ADE57469	Ade57469 Human Pro
16	1298.5	37.0	737	7	ADE57541	Ade57541 Human Pro
17	1298.5	37.0	737	7	ADE58354	Ade58354 Human Pro
18	1298.5	37.0	737	7	ADE57533	Ade57533 Human Pro
19	1298.5	37.0	737	7	ADE57537	Ade57537 Human Pro
20	1298.5	37.0	737	7	ADE57529	Ade57529 Human Pro
21	1296.5	37.0	737	6	ADA50073	Ada50073 Mouse pro
22	1296.5	37.0	737	7	AAE39167	Aae39167 Mouse pro
23	1290.5	36.8	737	7	ADE57531	Ade57531 Rat Prote
24	1290.5	36.8	737	7	ADE58352	Ade58352 Rat Prote
25	1290.5	36.8	737	7	ADE57535	Ade57535 Rat Prote
26	1290.5	36.8	737	7	ADE57539	Ade57539 Rat Prote
27	1290.5	36.8	737	7	ADE57527	Ade57527 Rat Prote
28	1278.5	36.5	739	4	ABB58410	Abb58410 Drosophil
29	1251	35.7	682	5	AAO18491	Aao18491 Human ins
30	1241.5	35.4	683	7	ADC37371	Adc37371 Nuclear f
31	1199.5	34.2	522	7	ADC37369	Adc37369 Nuclear f
32	1170	33.4	672	3	AAY93259	Aay93259 cDNA enco
33	1170	33.4	672	4	AAE13041	Aae13041 Human pro
34	1170	33.4	681	4	ABG06337	Abg06337 Novel hum
35	1169	33.3	672	2	AAR94765	Aar94765 Type III
36	1166	33.3	672	2	AAR66726	Aar66726 Protein k
37	1163	33.2	672	7	ADE65843	Ade65843 Human alp
38	1157.5	33.0	671	7	ADE56866	Ade56866 Rat Prote
39	1157.5	33.0	671	7	ADD45356	Add45356 Rat Prote
40	1157.5	33.0	671	7	ADD45358	Add45358 Human Pro
41	1157.5	33.0	673	6	AAO29576	Aao29576 Human Pc4
42	1157.5	33.0	673	7	ADE56868	Ade56868 Human Pro
43	1157.5	33.0	918	3	AAY70784	Aay70784 EGFP-PKCb
44	1155	32.9	672	5	ABB57302	Abb57302 Mouse isc
45	1155	32.9	916	2	AAW85023	Aaw85023 Amino aci

ALIGNMENTS

RESULT 1
ABG79705
ID ABG79705 standard; protein; 648 AA.
XX
AC ABG79705;
XX
DT 15-NOV-2002 (first entry)
XX
DE Tumour involved gene (TIG) splice variant protein, NV-36.
XX
KW Human; splice variant; tumour-involved gene; TIG;
KW pharmaceutical composition; cancer; diagnostic; tumour; gene therapy;
KW endothelial cell; cell differentiation; cell proliferation; apoptosis;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN US2002086384-A1.
XX
PD 04-JUL-2002.
XX
PF 13-MAR-2001; 2001US-00805020.
XX
PR 14-MAR-2000; 2000IL-00135402.
PR 16-MAY-2000; 2000IL-00136154.
XX
PA (LEVI/) LEVINE Z.
PA (DAVI/) DAVID A.
PA (ROMA/) ROMANO C.
PA (BERN/) BERNSTEIN J.
XX
PI Levine Z, David A, Romano C, Bernstein J;
XX
DR WPI; 2002-635679/68.
DR N-PSDB; ABS65235.
XX
PT Novel nucleic acid sequence, which is an alternative splicing variant of
PT tumor involved genes, useful for detecting cancer, predisposition to
PT cancer, for evaluating cancer state and in gene therapy for treating
PT cancer.
XX
PS Claim 4; Page 105-107; 180pp; English.
XX
CC The invention discloses isolated human nucleic acid alternative splicing
CC variants that are all tumour-involved genes (TIGs). The nucleic acids and
CC polypeptides are useful for determining the level of a nucleic acid or
CC polypeptide in a biological sample, for detecting a variant nucleic acid
CC or polypeptide sequence in a biological sample, for determining the level

CC of variant nucleic acid or polypeptide sequences in a biological sample
CC and for determining the ratio between the level of variant sequence in a
CC first biological sample and the level of the original sequence from which
CC the variant has been varied by alternative splicing in a second
CC biological sample and for raising antibodies. A pharmaceutical
CC composition comprising a carrier and the nucleic acid, is useful for
CC treating diseases (e.g. cancer) that can be ameliorated or cured by
CC increasing or decreasing the level of the encoded protein. The nucleic
CC acids are also useful for diagnostic purposes, especially for detecting
CC cancer or a predisposition to cancer, for evaluating the state or
CC aggressiveness of cancer disease, in basic research, for understanding
CC the physiological function of the original TIG, in targeting or
CC developing pharmaceuticals, for distinguishing various stages in the life
CC cycle of the same type of cells which may be helpful for the development
CC of pharmaceuticals for various cancer stages in which cell cycle is non-
CC normal, for determining mutations in tumour-involved genes and in gene
CC therapy. The polypeptides are useful for identifying compounds capable of
CC binding to the variant product and modulating its activity and for
CC modulating endothelial differentiation and proliferation, as well as to
CC modulate apoptosis either ex vivo or in vivo. The sequences presented in
CC ABG796700-ABG79705 are the new variants (NV) 1-36 proteins of the TIGs
CC disclosed
XX
SQ Sequence 648 AA;

Query Match 100.0%; Score 3506; DB 5; Length 648;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 648; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MSPFLRIGLSNFDCGSCQSCQGEAVNPYCAVLVKEYVESENGQMYIQKKPTMYPPWDSTF 60
Db    1 MSPFLRIGLSNFDCGSCQSCQGEAVNPYCAVLVKEYVESENGQMYIQKKPTMYPPWDSTF 60

Qy   61 DAHINKGRVMQIIVKGKNVDLISETTVELYSLAERCRKNNGKTEIWLELKPQGRMLMNAR 120
Db   61 DAHINKGRVMQIIVKGKNVDLISETTVELYSLAERCRKNNGKTEIWLELKPQGRMLMNAR 120

Qy  121 YFLEMSDTKDMNEFETEGFFALHQRRGAIKQAKVHHVKCHEFTATFFPQPTFCSVCHEFV 180
Db  121 YFLEMSDTKDMNEFETEGFFALHQRRGAIKQAKVHHVKCHEFTATFFPQPTFCSVCHEFV 180

Qy  181 WGLNKQGYQCRQCNAAIHKKCIDKVIAKCTGSAINSRETMFHKERFKIDMPHRFKVYNYK 240
Db  181 WGLNKQGYQCRQCNAAIHKKCIDKVIAKCTGSAINSRETMFHKERFKIDMPHRFKVYNYK 240

Qy  241 SPTFCEHCGTLLWGLARQGLKCDACGMNVHHRCQTKVANLCGINQKLMAEALAMIESTQQ 300
Db  241 SPTFCEHCGTLLWGLARQGLKCDACGMNVHHRCQTKVANLCGINQKLMAEALAMIESTQQ 300

Qy  301 ARCLRDTEQIFREGPVEIGLPCSIKNEARPPCLPTPGKREPQGISWESPLDEVDKMCHLP 360
Db  301 ARCLRDTEQIFREGPVEIGLPCSIKNEARPPCLPTPGKREPQGISWESPLDEVDKMCHLP 360

Qy  361 EPELNKERPSLQIKLKIEDFILHKMLGKGSFGKVFLAEFKKTNQFFAIKALKKDVVLMDD 420
Db  361 EPELNKERPSLQIKLKIEDFILHKMLGKGSFGKVFLAEFKKTNQFFAIKALKKDVVLMDD 420

Qy  421 DVECTMVEKRVLSLAWEHPFLTHMFCTFQTKENLFFVMEYLNGGDLMYHIQSCHKFDLSR 480
Db  421 DVECTMVEKRVLSLAWEHPFLTHMFCTFQTKENLFFVMEYLNGGDLMYHIQSCHKFDLSR 480

Qy  481 ATFYAAEIILGLQFLHSKGIVYRDLKLDNILLDKDGHIKIADFGMCKENMLGDAKTNTFC 540
Db  481 ATFYAAEIILGLQFLHSKGIVYRDLKLDNILLDKDGHIKIADFGMCKENMLGDAKTNTFC 540

Qy  541 GTPDYIAPEILLGQKYNHSVDWWSFGVLLYEMLIGQSPFHGQDEEELFHSIRMDNPFYPR 600
Db  541 GTPDYIAPEILLGQKYNHSVDWWSFGVLLYEMLIGQSPFHGQDEEELFHSIRMDNPFYPR 600

Qy  601 WLEKEAKDLLVKVRSEAKSVFIRRALGLLVSFLFLLVSNLHVANNDYY 648
Db  601 WLEKEAKDLLVKVRSEAKSVFIRRALGLLVSFLFLLVSNLHVANNDYY 648
```

RESULT 2
ABG79678
ID ABG79678 standard; protein; 615 AA.
XX
AC ABG79678;
XX
DT 15-NOV-2002 (first entry)
XX
DE Tumour involved gene (TIG) splice variant protein, NV-9.
XX
KW Human; splice variant; tumour-involved gene; TIG;
KW pharmaceutical composition; cancer; diagnostic; tumour; gene therapy;
KW endothelial cell; cell differentiation; cell proliferation; apoptosis;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN US2002086384-A1.
XX
PD 04-JUL-2002.
XX
PF 13-MAR-2001; 2001US-00805020.
XX
PR 14-MAR-2000; 2000IL-00135402.
PR 16-MAY-2000; 2000IL-00136154.
XX
PA (LEVI/) LEVINE Z.
PA (DAVI/) DAVID A.
PA (ROMA/) ROMANO C.
PA (BERN/) BERNSTEIN J.
XX
PI Levine Z, David A, Romano C, Bernstein J;
XX
DR WPI; 2002-635679/68.
DR N-PSDB; ABS65208.
XX
PT Novel nucleic acid sequence, which is an alternative splicing variant of
PT tumor involved genes, useful for detecting cancer, predisposition to
PT cancer, for evaluating cancer state and in gene therapy for treating
PT cancer.
XX
PS Claim 4; Page 73-75; 180pp; English.
XX
CC The invention discloses isolated human nucleic acid alternative splicing
CC variants that are all tumour-involved genes (TIGs). The nucleic acids and
CC polypeptides are useful for determining the level of a nucleic acid or
CC polypeptide in a biological sample, for detecting a variant nucleic acid
CC or polypeptide sequence in a biological sample, for determining the level
CC of variant nucleic acid or polypeptide sequences in a biological sample
CC and for determining the ratio between the level of variant sequence in a
CC first biological sample and the level of the original sequence from which
CC the variant has been varied by alternative splicing in a second
CC biological sample and for raising antibodies. A pharmaceutical
CC composition comprising a carrier and the nucleic acid, is useful for
CC treating diseases (e.g. cancer) that can be ameliorated or cured by
CC increasing or decreasing the level of the encoded protein. The nucleic
CC acids are also useful for diagnostic purposes, especially for detecting
CC cancer or a predisposition to cancer, for evaluating the state or
CC aggressiveness of cancer disease, in basic research, for understanding
CC the physiological function of the original TIG, in targeting or
CC developing pharmaceuticals, for distinguishing various stages in the life
CC cycle of the same type of cells which may be helpful for the development
CC of pharmaceuticals for various cancer stages in which cell cycle is non-
CC normal, for determining mutations in tumour-involved genes and in gene
CC therapy. The polypeptides are useful for identifying compounds capable of
CC binding to the variant product and modulating its activity and for
CC modulating endothelial differentiation and proliferation, as well as to
CC modulate apoptosis either ex vivo or in vivo. The sequences presented in
CC ABG796700-ABG79705 are the new variants (NV) 1-36 proteins of the TIGs
CC disclosed
XX
SQ Sequence 615 AA;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 16:38:50 ; Search time 952 Seconds
(without alignments)
10508.935 Million cell updates/sec

Title: US-09-805-020-36
Perfect score: 2355
Sequence: 1 gaattccgccagccccgcca.....ctccaaaacaataaagggga 2355

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2355	100.0	2355	6	ABS65235	Abs65235 cDNA enco
2	2331	99.0	2369	6	ABS65208	Abs65208 cDNA enco
3	1927.2	81.8	2754	4	AAF59810	Aaf59810 Human pro
4	1927.2	81.8	2754	6	ABL67085	Abl67085 Thyroid c
5	1927.2	81.8	2754	8	ACC85477	Acc85477 Human pro
6	1927.2	81.8	2754	9	ADE84903	Ade84903 Farnesyl
7	1857.4	78.9	2705	3	AAA35299	Aaa35299 Human ade
8	1857.4	78.9	2705	3	AAF21421	Aaf21421 Human low
9	1857.4	78.9	2705	5	AAH42213	Aah42213 Nucleotid
10	1857.4	78.9	2705	7	ABZ97115	Abz97115 Human nuc
11	1857.4	78.9	2705	7	ABZ22924	Abz22924 Human pro
12	1857.4	78.9	38644	3	AAA35302	Aaa35302 Human ade
13	1857.4	78.9	38644	3	AAF21424	Aaf21424 Human low
14	1857.4	78.9	38644	7	ABZ97118	Abz97118 Human nuc
15	1850.4	78.6	2698	5	AAH42214	Aah42214 Nucleotid
16	1836.4	78.0	2121	6	ABV78235	Abv78235 Human PKC
17	1836.4	78.0	2121	6	ABZ35811	Abz35811 Human PKC
18	1836.4	78.0	2121	6	ABX10054	Abx10054 Human PKC
19	1836.4	78.0	2121	6	ABL91776	Abl91776 Human pol
20	685.8	29.1	2909	2	AAQ01758	Aaq01758 cDNA sequ
21	685.8	29.1	2909	7	ABT41763	Abt41763 Toxicity
22	685.8	29.1	2909	9	ADB53627	Adb53627 Primary r
23	682.6	29.0	2891	2	AAQ57016	Aaq57016 PKC delta
24	670.2	28.5	2104	3	AAA35290	Aaa35290 Human ade
25	670.2	28.5	2104	3	AAF21412	Aaf21412 Human low
26	670.2	28.5	2104	4	AAC84222	Aac84222 Human pro
27	670.2	28.5	2104	7	ABZ97106	Abz97106 Human nuc
28	664.4	28.2	2031	6	ABV78229	Abv78229 Human PKC
29	664.4	28.2	2031	6	ABZ35805	Abz35805 Human PKC
30	664.4	28.2	2031	6	ABX10048	Abx10048 Human PKC
31	664.4	28.2	2031	6	ABL91770	Abl91770 Human pol
32	632.4	26.9	2163	3	AAA35289	Aaa35289 Human ade
33	632.4	26.9	2163	3	AAF21411	Aaf21411 Human low
34	632.4	26.9	2163	6	ABK83588	Abk83588 Human cDN
35	632.4	26.9	2163	7	ABZ97105	Abz97105 Human nuc
36	389	16.5	1423	3	AAA35292	Aaa35292 Human ade
37	389	16.5	1423	3	AAF21414	Aaf21414 Human low
38	389	16.5	1423	7	ABZ97108	Abz97108 Human nuc
39	376	16.0	2707	2	AAQ57014	Aaq57014 PKC epsil
40	374.4	15.9	2274	8	ADA50078	Ada50078 Protein k
41	374.4	15.9	2274	9	AAD59551	Aad59551 Mouse pro
42	374.2	15.9	2244	3	AAA35293	Aaa35293 Human ade
43	374.2	15.9	2244	3	AAF21415	Aaf21415 Human low
44	374.2	15.9	2244	6	AAL48609	Aal48609 Human ins
45	374.2	15.9	2244	7	ABZ97109	Abz97109 Human nuc

ALIGNMENTS

RESULT 1
ABS65235
ID ABS65235 standard; cDNA; 2355 BP.
XX
AC ABS65235;
XX
DT 15-NOV-2002 (first entry)
XX
DE cDNA encoding tumour involved gene (TIG) splice variant, NV-36.
XX
KW Human; ss; gene; splice variant; tumour-involved gene; TIG;
KW pharmaceutical composition; cancer; diagnostic; tumour; gene therapy;
KW endothelial cell; cell differentiation; cell proliferation; apoptosis;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN US2002086384-A1.
XX
PD 04-JUL-2002.
XX
PF 13-MAR-2001; 2001US-00805020.
XX
PR 14-MAR-2000; 2000IL-00135402.
PR 16-MAY-2000; 2000IL-00136154.
XX
PA (LEVI/) LEVINE Z.
PA (DAVI/) DAVID A.
PA (ROMA/) ROMANO C.
PA (BERN/) BERNSTEIN J.
XX
PI Levine Z, David A, Romano C, Bernstein J;
XX
DR WPI; 2002-635679/68.
DR P-PSDB; ABG79705.
XX
PT Novel nucleic acid sequence, which is an alternative splicing variant of
PT tumor involved genes, useful for detecting cancer, predisposition to
PT cancer, for evaluating cancer state and in gene therapy for treating
PT cancer.
XX
PS Claim 1; Page 64-65; 180pp; English.
XX
CC The invention discloses isolated human nucleic acid alternative splicing
CC variants that are all tumour-involved genes (TIGs). The nucleic acids and
CC polypeptides are useful for determining the level of a nucleic acid or

CC polypeptide in a biological sample, for detecting a variant nucleic acid
CC or polypeptide sequence in a biological sample, for determining the level
CC of variant nucleic acid or polypeptide sequences in a biological sample
CC and for determining the ratio between the level of variant sequence in a
CC first biological sample and the level of the original sequence from which
CC the variant has been varied by alternative splicing in a second
CC biological sample and for raising antibodies. A pharmaceutical
CC composition comprising a carrier and the nucleic acid, is useful for
CC treating diseases (e.g. cancer) that can be ameliorated or cured by
CC increasing or decreasing the level of the encoded protein. The nucleic
CC acids are also useful for diagnostic purposes, especially for detecting
CC cancer or a predisposition to cancer, for evaluating the state or
CC aggressiveness of cancer disease, in basic research, for understanding
CC the physiological function of the original TIG, in targeting or
CC developing pharmaceuticals, for distinguishing various stages in the life
CC cycle of the same type of cells which may be helpful for the development
CC of pharmaceuticals for various cancer stages in which cell cycle is non-
CC normal, for determining mutations in tumour-involved genes and in gene
CC therapy. The polypeptides are useful for identifying compounds capable of
CC binding to the variant product and modulating its activity and for
CC modulating endothelial differentiation and proliferation, as well as to
CC modulate apoptosis either ex vivo or in vivo. The sequences presented in
CC ABS652200-ABS65235 are the coding sequences for the new variants (NV) 1-
CC 36 of the TIGs disclosed
XX
SQ Sequence 2355 BP; 667 A; 537 C; 594 G; 557 T; 0 U; 0 Other;

Query Match 100.0%; Score 2355; DB 6; Length 2355;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GAATTCCGCCAGCCCCGCCAGTCCCCGCGCAGTCCCCGCGCAGTCCCAGCGCCACCGGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAATTCCGCCAGCCCCGCCAGTCCCCGCGCAGTCCCCGCGCAGTCCCAGCGCCACCGGGC 60

Qy         61 AGCAGCGGCGCCGTGCTCGCTCCAGGGCGCAACCATGTCGCCATTTCTTCGGATTGGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCAGCGGCGCCGTGCTCGCTCCAGGGCGCAACCATGTCGCCATTTCTTCGGATTGGCTT 120

Qy        121 GTCCAACTTTGACTGCGGGTCCTGCCAGTCTTGTCAGGGCGAGGCTGTTAACCCTTACTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCCAACTTTGACTGCGGGTCCTGCCAGTCTTGTCAGGGCGAGGCTGTTAACCCTTACTG 180

Qy        181 TGCTGTGCTCGTCAAAGAGTATGTCGAATCAGAGAACGGGCAGATGTATATCCAGAAAAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTGTGCTCGTCAAAGAGTATGTCGAATCAGAGAACGGGCAGATGTATATCCAGAAAAA 240

Qy        241 GCCTACCATGTACCCACCCTGGGACAGCACTTTTGATGCCCATATCAACAAGGGAAGAGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCTACCATGTACCCACCCTGGGACAGCACTTTTGATGCCCATATCAACAAGGGAAGAGT 300

Qy        301 CATGCAGATCATTGTGAAAGGCAAAAACGTGGACCTCATCTCTGAAACCACCGTGGAGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATGCAGATCATTGTGAAAGGCAAAAACGTGGACCTCATCTCTGAAACCACCGTGGAGCT 360

Qy        361 CTACTCGCTGGCTGAGAGGTGCAGGAAGAACAACGGGAAGACAGAAATATGGTTAGAGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTACTCGCTGGCTGAGAGGTGCAGGAAGAACAACGGGAAGACAGAAATATGGTTAGAGCT 420

Qy        421 GAAACCTCAAGGCCGAATGCTAATGAATGCAAGATACTTTCTGGAAATGAGTGACACAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAACCTCAAGGCCGAATGCTAATGAATGCAAGATACTTTCTGGAAATGAGTGACACAAA 480

Qy        481 GGACATGAATGAATTTGAGACGGAAGGCTTCTTTGCTTTGCATCAGCGCCGGGGTGCCAT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGACATGAATGAATTTGAGACGGAAGGCTTCTTTGCTTTGCATCAGCGCCGGGGTGCCAT 540

Qy        541 CAAGCAGGCAAAGGTCCACCACGTCAAGTGCCACGAGTTCACTGCCACCTTCTTCCCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCAGGCAAAGGTCCACCACGTCAAGTGCCACGAGTTCACTGCCACCTTCTTCCCACA 600

Qy        601 GCCCACATTTTGCTCTGTCTGCCACGAGTTTGTCTGGGGCCTGAACAAACAGGGCTACCA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCCCACATTTTGCTCTGTCTGCCACGAGTTTGTCTGGGGCCTGAACAAACAGGGCTACCA 660

Qy        661 GTGCCGACAATGCAATGCAGCAATTCACAAGAAGTGTATTGATAAAGTTATAGCAAAGTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTGCCGACAATGCAATGCAGCAATTCACAAGAAGTGTATTGATAAAGTTATAGCAAAGTG 720

Qy        721 CACAGGATCAGCTATCAATAGCCGAGAAACCATGTTCCACAAGGAGAGATTCAAAATTGA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CACAGGATCAGCTATCAATAGCCGAGAAACCATGTTCCACAAGGAGAGATTCAAAATTGA 780

Qy        781 CATGCCACACAGATTTAAAGTCTACAATTACAAGAGCCCGACCTTCTGTGAACACTGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGCCACACAGATTTAAAGTCTACAATTACAAGAGCCCGACCTTCTGTGAACACTGTGG 840

Qy        841 GACCCTGCTGTGGGGACTGGCACGGCAAGGACTCAAGTGTGATGCATGTGGCATGAATGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GACCCTGCTGTGGGGACTGGCACGGCAAGGACTCAAGTGTGATGCATGTGGCATGAATGT 900

Qy        901 GCATCATAGATGCCAGACAAAGGTGGCCAACCTTTGTGGCATAAACCAGAAGCTAATGGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCATCATAGATGCCAGACAAAGGTGGCCAACCTTTGTGGCATAAACCAGAAGCTAATGGC 960

Qy        961 TGAAGCGCTGGCCATGATTGAGAGCACTCAACAGGCTCGCTGCTTAAGAGATACTGAACA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGAAGCGCTGGCCATGATTGAGAGCACTCAACAGGCTCGCTGCTTAAGAGATACTGAACA 1020

Qy       1021 GATCTTCAGAGAAGGTCCGGTTGAAATTGGTCTCCCATGCTCCATCAAAAATGAAGCAAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GATCTTCAGAGAAGGTCCGGTTGAAATTGGTCTCCCATGCTCCATCAAAAATGAAGCAAG 1080

Qy       1081 GCCGCCATGTTTACCGACACCGGGAAAAAGAGAGCCTCAGGGCATTTCCTGGGAGTCTCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCGCCATGTTTACCGACACCGGGAAAAAGAGAGCCTCAGGGCATTTCCTGGGAGTCTCC 1140

Qy       1141 GTTGGATGAGGTGGATAAAATGTGCCATCTTCCAGAACCTGAACTGAACAAAGAAAGACC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTTGGATGAGGTGGATAAAATGTGCCATCTTCCAGAACCTGAACTGAACAAAGAAAGACC 1200

Qy       1201 ATCTCTGCAGATTAAACTAAAAATTGAGGATTTTATCTTGCACAAAATGTTGGGGAAAGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATCTCTGCAGATTAAACTAAAAATTGAGGATTTTATCTTGCACAAAATGTTGGGGAAAGG 1260

Qy       1261 AAGTTTTGGCAAGGTCTTCCTGGCAGAATTCAAGAAAACCAATCAATTTTTCGCAATAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAGTTTTGGCAAGGTCTTCCTGGCAGAATTCAAGAAAACCAATCAATTTTTCGCAATAAA 1320

Qy       1321 GGCCTTAAAGAAAGATGTGGTCTTGATGGACGATGATGTTGAGTGCACGATGGTAGAGAA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCCTTAAAGAAAGATGTGGTCTTGATGGACGATGATGTTGAGTGCACGATGGTAGAGAA 1380

Qy       1381 GAGAGTTCTTTCCTTGGCCTGGGAGCATCCGTTTCTGACGCACATGTTTTGTACATTCCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GAGAGTTCTTTCCTTGGCCTGGGAGCATCCGTTTCTGACGCACATGTTTTGTACATTCCA 1440

Qy       1441 GACCAAGGAAAACCTCTTTTTTGTGATGGAGTACCTCAACGGAGGGGACTTAATGTACCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GACCAAGGAAAACCTCTTTTTTGTGATGGAGTACCTCAACGGAGGGGACTTAATGTACCA 1500

Qy       1501 CATCCAAAGCTGCCACAAGTTCGACCTTTCCAGAGCGACGTTTTATGCTGCTGAAATCAT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CATCCAAAGCTGCCACAAGTTCGACCTTTCCAGAGCGACGTTTTATGCTGCTGAAATCAT 1560

Qy       1561 TCTTGGTCTGCAGTTCCTTCATTCCAAAGGAATAGTCTACAGGGACCTGAAGCTAGATAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCTTGGTCTGCAGTTCCTTCATTCCAAAGGAATAGTCTACAGGGACCTGAAGCTAGATAA 1620

Qy       1621 CATCCTGTTAGACAAAGATGGACATATCAAGATCGCGGATTTTGGAATGTGCAAGGAGAA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CATCCTGTTAGACAAAGATGGACATATCAAGATCGCGGATTTTGGAATGTGCAAGGAGAA 1680

Qy       1681 CATGTTAGGAGATGCCAAGACGAATACCTTCTGTGGGACACCTGACTACATCGCCCCAGA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CATGTTAGGAGATGCCAAGACGAATACCTTCTGTGGGACACCTGACTACATCGCCCCAGA 1740
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 17:56:26 ; Search time 9050 Seconds
(without alignments)
11278.767 Million cell updates/sec

Title: US-09-805-020-36
Perfect score: 2355
Sequence: 1 gaattccgccagccccgcca.....ctccaaaacaataaagggga 2355

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1930.4	82.0	3292	6	AX780110	AX780110 Sequence
	2	1927.2	81.8	2754	6	AR130752	AR130752 Sequence
	3	1927.2	81.8	2754	6	AX334913	AX334913 Sequence
	4	1927.2	81.8	2754	6	AX774806	AX774806 Sequence
	5	1927.2	81.8	2754	9	HUMPKC	L01087 Human prote
	6	1857.4	78.9	2705	9	HUMPKCTH	L07032 Human prote
	7	1836.4	78.0	2121	6	AX481505	AX481505 Sequence
	8	1834.8	77.9	3370	9	HSM807803	BX647657 Homo sapi
	9	1474.6	62.6	3313	10	MUSPKCT	D11091 Mouse mRNA
	10	813.8	34.6	2184	10	AB062122	AB062122 Mus muscu
	11	743.6	31.6	2977	10	AF219629	AF219629 Rattus sp
	12	712	30.2	2646	10	AB011812	AB011812 Mus muscu
	13	691	29.3	2538	10	MMNPKCD	X60304 M.musculus
	14	690	29.3	2564	10	MUSPKCDA	M69042 Mouse prote
	15	686.2	29.1	2025	10	AF251036	AF251036 Mus muscu
	16	685.8	29.1	2909	6	E02147	E02147 Rat protein
	17	685.8	29.1	2909	10	RATPKCDA	M18330 Rat protein
	18	684.2	29.1	2891	6	A37237	A37237 Sequence 6
	19	680.2	28.9	2069	5	AB109739	AB109739 Xenopus l
	20	677.4	28.8	2100	5	AB109740	AB109740 Xenopus l
	21	673.6	28.6	2693	9	BC043350	BC043350 Homo sapi
	22	671.8	28.5	2104	9	HUMPKCD13Y	L07861 Human prote
	23	670.2	28.5	2104	6	AR153554	AR153554 Sequence
	24	670.2	28.5	2104	6	BD262869	BD262869 Antisense
	25	670.2	28.5	2104	6	AR380646	AR380646 Sequence
	26	670.2	28.5	2104	6	AX771580	AX771580 Sequence
	27	670.2	28.5	2104	6	AX779975	AX779975 Sequence
	28	670.2	28.5	2104	9	HUMPCKD13X	L07860 Human prote
	29	664.4	28.2	2031	6	AX481499	AX481499 Sequence
	30	664.4	28.2	2031	12	AY335687	AY335687 Synthetic
	31	634.2	26.9	2918	5	BC049327	BC049327 Danio rer
	32	632.4	26.9	2163	9	HUMPKSCD	D10495 Homo sapien
c	33	592	25.1	189767	9	AL137145	AL137145 Human DNA
c	34	590.4	25.1	137072	2	AL353893	AL353893 Homo sapi
	35	532	22.6	1848	9	AK130150	AK130150 Homo sapi
	36	498.2	21.2	2517	10	BC051416	BC051416 Mus muscu
	37	389.4	16.5	3455	3	SRNPKCSR	Y13103 S.raphanus
	38	389	16.5	1423	9	HSPRKNCDA	Z22521 H.sapiens p
	39	376	16.0	2235	10	AF028009	AF028009 Mus muscu
	40	376	16.0	2707	6	A37235	A37235 Sequence 4
	41	374.4	15.9	2214	10	AF325507	AF325507 Mus muscu
	42	374.4	15.9	2274	6	AR282728	AR282728 Sequence
	43	374.2	15.9	2244	9	HSPKCE	X65293 H.sapiens m
	44	372.8	15.8	2704	6	E02148	E02148 Rat protein
	45	372.8	15.8	2704	10	RATPKCEA	M18331 Rat protein

ALIGNMENTS

RESULT 1
AX780110
LOCUS AX780110 3292 bp DNA linear PAT 14-JUL-2003
DEFINITION Sequence 2267 from Patent WO03039443.
ACCESSION AX780110
VERSION AX780110.1 GI:32697104
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,
Dugas,M., Eils,R., Brors,B. and Mergenthaler,S.
TITLE Novel genetic markers for leukemias

```
  JOURNAL   Patent: WO 03039443-A 2267 15-MAY-2003;
            Deutsches Krebsforschungszentrum (DE) ;
            Ludwig-Maximilian-Universitaet Muenchen (DE) ; Haferlach, Torsten,
            PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)
FEATURES             Location/Qualifiers
     source          1. .3292
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             82.0%;  Score 1930.4;  DB 6;  Length 3292;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1931;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GAATTCCGCCAGCCCCGCCAGTCCCCGCGCAGTCCCCGCGCAGTCCCAGCGCCACCGGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAATTCCGCCAGCCCCGCCAGTCCCCGCGCAGTCCCCGCGCAGTCCCAGCGCCACCGGGC 60

Qy         61 AGCAGCGGCGCCGTGCTCGCTCCAGGGCGCAACCATGTCGCCATTTCTTCGGATTGGCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCAGCGGCGCCGTGCTCGCTCCAGGGCGCAACCATGTCGCCATTTCTTCGGATTGGCTT 120

Qy        121 GTCCAACTTTGACTGCGGGTCCTGCCAGTCTTGTCAGGGCGAGGCTGTTAACCCTTACTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCCAACTTTGACTGCGGGTCCTGCCAGTCTTGTCAGGGCGAGGCTGTTAACCCTTACTG 180

Qy        181 TGCTGTGCTCGTCAAAGAGTATGTCGAATCAGAGAACGGGCAGATGTATATCCAGAAAAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTGTGCTCGTCAAAGAGTATGTCGAATCAGAGAACGGGCAGATGTATATCCAGAAAAA 240

Qy        241 GCCTACCATGTACCCACCCTGGGACAGCACTTTTGATGCCCATATCAACAAGGGAAGAGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCTACCATGTACCCACCCTGGGACAGCACTTTTGATGCCCATATCAACAAGGGAAGAGT 300

Qy        301 CATGCAGATCATTGTGAAAGGCAAAAACGTGGACCTCATCTCTGAAACCACCGTGGAGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATGCAGATCATTGTGAAAGGCAAAAACGTGGACCTCATCTCTGAAACCACCGTGGAGCT 360

Qy        361 CTACTCGCTGGCTGAGAGGTGCAGGAAGAACAACGGGAAGACAGAAATATGGTTAGAGCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTACTCGCTGGCTGAGAGGTGCAGGAAGAACAACGGGAAGACAGAAATATGGTTAGAGCT 420

Qy        421 GAAACCTCAAGGCCGAATGCTAATGAATGCAAGATACTTTCTGGAAATGAGTGACACAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAACCTCAAGGCCGAATGCTAATGAATGCAAGATACTTTCTGGAAATGAGTGACACAAA 480

Qy        481 GGACATGAATGAATTTGAGACGGAAGGCTTCTTTGCTTTGCATCAGCGCCGGGGTGCCAT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGACATGAATGAATTTGAGACGGAAGGCTTCTTTGCTTTGCATCAGCGCCGGGGTGCCAT 540

Qy        541 CAAGCAGGCAAAGGTCCACCACGTCAAGTGCCACGAGTTCACTGCCACCTTCTTCCCACA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCAGGCAAAGGTCCACCACGTCAAGTGCCACGAGTTCACTGCCACCTTCTTCCCACA 600

Qy        601 GCCCACATTTTGCTCTGTCTGCCACGAGTTTGTCTGGGGCCTGAACAAACAGGGCTACCA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCCCACATTTTGCTCTGTCTGCCACGAGTTTGTCTGGGGCCTGAACAAACAGGGCTACCA 660

Qy        661 GTGCCGACAATGCAATGCAGCAATTCACAAGAAGTGTATTGATAAAGTTATAGCAAAGTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTGCCGACAATGCAATGCAGCAATTCACAAGAAGTGTATTGATAAAGTTATAGCAAAGTG 720

Qy        721 CACAGGATCAGCTATCAATAGCCGAGAAACCATGTTCCACAAGGAGAGATTCAAAATTGA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CACAGGATCAGCTATCAATAGCCGAGAAACCATGTTCCACAAGGAGAGATTCAAAATTGA 780

Qy        781 CATGCCACACAGATTTAAAGTCTACAATTACAAGAGCCCGACCTTCTGTGAACACTGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGCCACACAGATTTAAAGTCTACAATTACAAGAGCCCGACCTTCTGTGAACACTGTGG 840

Qy        841 GACCCTGCTGTGGGGACTGGCACGGCAAGGACTCAAGTGTGATGCATGTGGCATGAATGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        841 GACCCTGCTGTGGGGACTGGCACGGCAAGGACTCAAGTGTGATGCATGTGGCATGAATGT 900

Qy        901 GCATCATAGATGCCAGACAAAGGTGGCCAACCTTTGTGGCATAAACCAGAAGCTAATGGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCATCATAGATGCCAGACAAAGGTGGCCAACCTTTGTGGCATAAACCAGAAGCTAATGGC 960

Qy        961 TGAAGCGCTGGCCATGATTGAGAGCACTCAACAGGCTCGCTGCTTAAGAGATACTGAACA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGAAGCGCTGGCCATGATTGAGAGCACTCAACAGGCTCGCTGCTTAAGAGATACTGAACA 1020

Qy       1021 GATCTTCAGAGAAGGTCCGGTTGAAATTGGTCTCCCATGCTCCATCAAAAATGAAGCAAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GATCTTCAGAGAAGGTCCGGTTGAAATTGGTCTCCCATGCTCCATCAAAAATGAAGCAAG 1080

Qy       1081 GCCGCCATGTTTACCGACACCGGGAAAAAGAGAGCCTCAGGGCATTTCCTGGGAGTCTCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCGCCATGTTTACCGACACCGGGAAAAAGAGAGCCTCAGGGCATTTCCTGGGAGTCTCC 1140

Qy       1141 GTTGGATGAGGTGGATAAAATGTGCCATCTTCCAGAACCTGAACTGAACAAAGAAAGACC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTTGGATGAGGTGGATAAAATGTGCCATCTTCCAGAACCTGAACTGAACAAAGAAAGACC 1200

Qy       1201 ATCTCTGCAGATTAAACTAAAAATTGAGGATTTTATCTTGCACAAAATGTTGGGGAAAGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATCTCTGCAGATTAAACTAAAAATTGAGGATTTTATCTTGCACAAAATGTTGGGGAAAGG 1260

Qy       1261 AAGTTTTGGCAAGGTCTTCCTGGCAGAATTCAAGAAAACCAATCAATTTTTCGCAATAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAGTTTTGGCAAGGTCTTCCTGGCAGAATTCAAGAAAACCAATCAATTTTTCGCAATAAA 1320

Qy       1321 GGCCTTAAAGAAAGATGTGGTCTTGATGGACGATGATGTTGAGTGCACGATGGTAGAGAA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCCTTAAAGAAAGATGTGGTCTTGATGGACGATGATGTTGAGTGCACGATGGTAGAGAA 1380

Qy       1381 GAGAGTTCTTTCCTTGGCCTGGGAGCATCCGTTTCTGACGCACATGTTTTGTACATTCCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GAGAGTTCTTTCCTTGGCCTGGGAGCATCCGTTTCTGACGCACATGTTTTGTACATTCCA 1440

Qy       1441 GACCAAGGAAAACCTCTTTTTTGTGATGGAGTACCTCAACGGAGGGGACTTAATGTACCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GACCAAGGAAAACCTCTTTTTTGTGATGGAGTACCTCAACGGAGGGGACTTAATGTACCA 1500

Qy       1501 CATCCAAAGCTGCCACAAGTTCGACCTTTCCAGAGCGACGTTTTATGCTGCTGAAATCAT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CATCCAAAGCTGCCACAAGTTCGACCTTTCCAGAGCGACGTTTTATGCTGCTGAAATCAT 1560

Qy       1561 TCTTGGTCTGCAGTTCCTTCATTCCAAAGGAATAGTCTACAGGGACCTGAAGCTAGATAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCTTGGTCTGCAGTTCCTTCATTCCAAAGGAATAGTCTACAGGGACCTGAAGCTAGATAA 1620

Qy       1621 CATCCTGTTAGACAAAGATGGACATATCAAGATCGCGGATTTTGGAATGTGCAAGGAGAA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CATCCTGTTAGACAAAGATGGACATATCAAGATCGCGGATTTTGGAATGTGCAAGGAGAA 1680

Qy       1681 CATGTTAGGAGATGCCAAGACGAATACCTTCTGTGGGACACCTGACTACATCGCCCCAGA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CATGTTAGGAGATGCCAAGACGAATACCTTCTGTGGGACACCTGACTACATCGCCCCAGA 1740

Qy       1741 GATCTTGCTGGGTCAGAAATACAACCACTCTGTGGACTGGTGGTCCTTCGGGGTTCTCCT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GATCTTGCTGGGTCAGAAATACAACCACTCTGTGGACTGGTGGTCCTTCGGGGTTCTCCT 1800

Qy       1801 TTATGAAATGCTGATTGGTCAGTCGCCTTTCCACGGGCAGGATGAGGAGGAGCTCTTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TTATGAAATGCTGATTGGTCAGTCGCCTTTCCACGGGCAGGATGAGGAGGAGCTCTTCCA 1860

Qy       1861 CTCCATCCGCATGGACAATCCCTTTTACCCACGGTGGCTGGAGAAGGAAGCAAAGGACCT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTCCATCCGCATGGACAATCCCTTTTACCCACGGTGGCTGGAGAAGGAAGCAAAGGACCT 1920

Qy       1921 TCTGGTGAAGGT 1932
              |||||||||| |
Db       1921 TCTGGTGAAGCT 1932
```